

# Open Research Online

---

The Open University's repository of research publications and other research outputs

## A quasi-current representation for information needs inspired by Two-State Vector Formalism

### Journal Item

#### How to cite:

Wang, Panpan; Hou, Yuexian; Li, Jingfei; Zhang, Yazhou; Song, Dawei and Li, Wenjie (2017). A quasi-current representation for information needs inspired by Two-State Vector Formalism. *Physica A: Statistical Mechanics and its Applications*, 482 pp. 627–637.

For guidance on citations see [FAQs](#).

© 2017 Elsevier B.V.



<https://creativecommons.org/licenses/by-nc-nd/4.0/>

Version: Accepted Manuscript

Link(s) to article on publisher's website:

<http://dx.doi.org/doi:10.1016/j.physa.2017.04.145>

---

Copyright and Moral Rights for the articles on this site are retained by the individual authors and/or other copyright owners. For more information on Open Research Online's data [policy](#) on reuse of materials please consult the policies page.

---

[oro.open.ac.uk](http://oro.open.ac.uk)

## Accepted Manuscript

A quasi-current representation for information needs inspired by  
Two-State Vector Formalism

Panpan Wang, Yuexian Hou, Jingfei Li, Yazhou Zhang, Dawei Song,  
Wenjie Li

PII: S0378-4371(17)30455-7

DOI: <http://dx.doi.org/10.1016/j.physa.2017.04.145>

Reference: PHYSA 18247

To appear in: *Physica A*

Received date: 8 October 2016

Revised date: 20 April 2017

Please cite this article as: P. Wang, Y. Hou, J. Li, Y. Zhang, D. Song, W. Li, A quasi-current representation for information needs inspired by Two-State Vector Formalism, *Physica A* (2017), <http://dx.doi.org/10.1016/j.physa.2017.04.145>

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.



Highlights

1. A novel Quantum-based Information Retrieval (QIR) Model inspired by Two-State Vector Formalism is proposed.
2. The current query and the previous query in a search session are applied to construct a quasi-current representation for users' information needs.
3. Extensive experiments have shown its effectiveness.

# A Quasi-Current Representation for Information Needs Inspired by Two-State Vector Formalism

Panpan Wang<sup>a</sup>, Yuexian Hou<sup>a,\*</sup>, Jingfei Li<sup>a</sup>, Yazhou Zhang<sup>a</sup>, Dawei Song<sup>a,b</sup>,  
Wenjie Li<sup>c</sup>

<sup>a</sup>*Tianjin Key Laboratory of Cognitive Computing and Application, School of Computer  
Science and Technology, Tianjin University, 300350, No.135 Yaguan Road, Jinnan  
District, Tianjin, P.R.China*

<sup>b</sup>*The Computing and Communications Department, the Open University, Walton Hall,  
Milton Keynes, MK7 6AA. United Kingdom*

<sup>c</sup>*Department of Computing and Communications, The Open University, UK*

---

## Abstract

Recently, a number of quantum theory (QT)-based information retrieval (IR) models have been proposed for modeling session search task that users issue queries continuously in order to describe their evolving information needs (IN). However, the standard formalism of QT cannot provide a complete description for users' current IN in a sense that it does not take the 'future' information into consideration. Therefore, to seek a more proper and complete representation for users' IN, we construct a representation of quasi-current IN inspired by an emerging Two-State Vector Formalism (TSVF). With the enlightenment of the completeness of TSVF, a "two-state vector" derived from the 'future' (the current query) and the 'history' (the previous query) is employed to describe users' quasi-current IN in a more complete way. Extensive experiments are conducted on the session tracks of TREC 2013 & 2014, and show that our model outperforms a series of compared IR models.

**Keywords:** Information Retrieval, Two-State Vector Formalism, Quantum Theory, Session Search

---

## 1. Introduction

Recently, quantum theory (QT), as an important formalism for modeling Information Retrieval (IR) tasks, has attracted increasing attention. van Rijs-

---

\*Corresponding authors: Yuexian Hou (yxhou@tju.edu.cn)

bergen, in his seminal book [1], proposed for the first time to employ QT as  
 5 a unified theoretical formalism for modeling IR tasks. The book showed that  
 major IR models (e.g., logical, probabilistic and vector) can be subsumed by a  
 single mathematical formalism in Hilbert vector space (which can be a complex  
 space). In that book, some notions in IR are translated into analogous notions  
 in QT, such as mapping a document into a **state vector**, regarding each docu-  
 10 ment as a **superposition** of words, and replacing the cosine correlation between  
 query and documents with **inner product**. Beyond that, QT can help address  
 some problems for IR tasks [2, 3, 4].

Following the pioneering work, a series of Quantum Theory-based IR (QIR)  
 models were developed and motivated by quantum probability framework. A  
 15 representative was Quantum Language Model (QLM) [5], which was presented  
 to model term dependencies in IR and gained good performance for ad-hoc re-  
 trieval tasks. However, QLM was solely targeted on single ad-hoc queries and  
 limited its further application on the dynamic search tasks, e.g., session search.  
 To solve that issue, Li et al. [6] developed an adaptive contextual QLM which u-  
 20 tilized a density matrix transformation framework to capture the dynamic infor-  
 mation (historical queries and clicked documents) within users' search sessions.  
 Then a session-based QLM [7] was also put forward to divide those interaction  
 information into positive and negative feedback to model the evolution of users'  
 information needs (IN). Later, with the inspiration of "quantum interference",  
 25 the interactive information in a session was used to construct a new superposed  
 state of a document in the IN space [8]. All these models mainly focused on uti-  
 lizing some concepts or phenomena of QT to describe users' IN. However, they  
 ignore an important fact that the dynamic evolution of users' IN in a session  
 search is supposed be a Markovian-like process from the perspective of QT.

30 Under the standard formalism of QT, the evolution of a quantum system  
 is a Markovian-like process [9] in the sense that the current quantum state of a  
 system can be completely described by the result of the last measurement. In  
 this paper, a user's IN is analogous to the quantum state, and a series of issued  
 queries are considered as a series of measurement results. Hence, if we use the  
 35 standard formalism of QT to model the evolution of a user's IN, it should also  
 be a Markovian-like process, in which the current quantum state is determined  
 by the last query (also called the current query in this paper), and is irrelevant  
 with the earlier queries. Note that the above observation could only make sense

in a QT framework. We do not suppose that it is universally valid.

40 However, it is still argued that the standard QT cannot completely characterize a quantum state in a sense that it does not take the information from the ‘future’ into consideration [10]. In short, the current query might be a proper description for the state of users’ current IN, but it is not complete. Note that the ‘complete quantum description’ corresponds to the completeness under A-  
45 haronov, Bergmann, and Lebowitz (ABL) principle mentioned in [9]. To seek a proper representation for users’ IN, we construct a quasi-current IN, which contains more complete information than the real current IN, inspired by an emerging Two-State Vector Formalism (TSVF).

Different from the standard formalism of QT, TSVF equips a time-symmetric  
50 formulation for QT, which the current system is described by a two-state vector that contains a backward-evolving vector (named post-selected vector) defined by the results of measurement performed on this system in future and a forward-evolving vector (named pre-selected vector). Some evidences have pointed out that such a formalism can provide more complete information than standard QT  
55 [11, 12, 13]. For example, the result of a measurement of  $\sigma_x, \sigma_y, \sigma_z$  performed on a spin- $\frac{1}{2}$  particle at a given time cannot be inferred under the standard QT prescriptions due to the three non-commuting measurements. However, those results can be ascertained with probability 1 under TSVF framework [14]. Currently, TSVF is mainly applied to make interpretations for some unusual quantum phenomena [15, 16], such as, Three-boxes paradox and Cheshire Cat. In  
60 this article, we apply TSVF to IR tasks for the first time and further investigate how much necessary information we should use to represent the quasi-current IN for QIR models.

According to TSVF, it is incomplete to model users’ current IN only by  
65 the current query (the query that is not retrieved by search engine yet). To construct a more complete description, our quasi-current IN should be modeled by the historical IN and the future IN. We choose the previous query to represent the historical IN since the earlier queries and interactions in search session are actually deviated from the real IN in accordance with the quantum philosophy.  
70 Taking session 2 in 2013 Session Track as an example, the current query “where to buy scooters” tells us that users’ current information needs are finding a place to buy scooters. And the previous query “scooter stores” expresses similar meanings. However, the third last query is “scooter price” which is obviously

another subtopic about scooter. If the current query and the previous query  
75 are combined together to represent users' current IN, the weight of these key  
words will be increased during retrieving the relevant documents. If the third  
last query is added, some documents containing the word 'price' will be also  
retrieved and further disturb user searching for right answers. Therefore, the  
current query and the previous query combined together can provide a more  
80 proper and complete description for users' current IN. Note that the amount of  
information in our work is not a statistical or information-theoretical concept,  
but rather an empirical consideration about a description of the past and future  
information needs of the user. To this end, we formalized a new target function  
for original QLM [5] inspired by TSVF and obtained a reliable representation  
85 for current information needs as density matrices by maximizing the formalized  
target function. Due to the maximum likelihood estimation methods in QLM  
cannot ensure global convergence, we replace the original training algorithm  
with *DilutedRpR* [17]. Extensive experiments are conducted on the session  
tracks of TREC 2013 and TREC 2014, which show that our model outperforms  
90 a series of compared IR models.

## 2. Related Work

van Rijsbergen (2004) provided a mathematical framework based on quan-  
tum theory (QT) for the foundations of IR [1]. Using this framework, a doc-  
ument could be represented as a vector in Hilbert space, and the document's  
95 relevance could be described by a Hermitian operator. Except that all the usual  
QT notions, such as superposition state and observation, had their IR-theoretic  
analogues, the standard QT can be applied to address problems in IR, such as  
pseudo-relevance feedback and relevance feedback.

Succeedingly, a series of quantum-based IR (QIR) models were proposed  
100 under the mathematical framework of QT, and especially motivated by some  
quantum phenomena, such as "quantum interference" (QI) and "photon po-  
larization" [4, 18, 19, 20, 21]. Zuccon and Azzopardi [4] proposed a Quantum  
Probability Ranking Principle (QPRP) that implicitly captured dependencies  
between documents through QI. Such a quantum effect was also used for model-  
105 ing interactions between latent topics [19]. In order to perform query expansion,  
a Quantum Entropy Minimization (QEM) model was proposed for learning se-  
mantic representations for words and phrases [22]. Later, Zhang [8] expanded

QT to session search by constructing a new superposition state of each document in the information needs space and incorporating QI in query expansion. Except for QI, other quantum notions were applied to address IR problems. In order to alleviate the query-drift problem caused by expanded query, Zhang [18] derived a novel fusion approach inspired by photon polarization. Then, Zhao [3] developed a re-ranking approach explored by another important QT concept, namely “quantum measurement”. Besides, Piwowarski et al.[2] presented to capture different aspects of information needs using tensor space, and achieved acceptable performance in an ad-hoc retrieval task.

Although those QIR models mentioned above have achieved good results and made heuristic utilization of the quantum concepts, they did not give a clear interpretation about quantum probability. A Quantum Language Model (QLM) [5] was proposed to model term dependencies in IR and gained good performance for ad-hoc retrieval. Subsequently, a series of variants [6, 7, 23] are proposed based on the QLM in order to make an expansion in wider IR scenarios. For example, an adaptive contextual QLM model was developed [6] to model users’ dynamic IN in the context of users interaction. Our model is also proposed based on QLM and employed for session search task, but it is developed from a totally different perspective. In contrast with [6] which utilizes as much session information as possible to enhance the retrieval performance, we focus on finding necessary information to represent users’ current IN from existing session data inspired by Two-State Vector Formalism (TSVF). In addition, we choose another different maximum likelihood estimation (MLE) method which can ensure the global convergence and obtain a Hermitian density matrix [17], see more detail in Section 4.1 and 4.2.

### 3. TSVF and Its Analogy to IR

In this section, we will first introduce the basic knowledge of quantum theory. Then, the difference between the standard formalism of QT and TSVF will be presented. Finally, we will describe the adoption of TSVF framework in session search task.

#### 3.1. Preliminary of Quantum Theory

In QT, the quantum probability space is naturally encapsulated in an infinite Hilbert space, noted as  $\mathbb{H}^n$ , which is an abstract vector space processing



the structure of the inner product. A finite dimensional space is sufficient for the work reported in this paper, thus, we limit our researches to a finite real space, denoted as  $\mathbb{R}^n$ . With the Dirac's notation, a quantum state  $\vec{u} \in \mathbb{R}^n$  and its transpose  $\vec{u}^T$  are respectively expressed as a ket  $|u\rangle$  and a bra  $\langle u|$ . Supposing  $|e_1\rangle, |e_2\rangle, \dots, |e_n\rangle$  forms an orthogonal basis for  $\mathbb{R}^n$ , then each unit vector  $|v\rangle$  can be uniquely written as the superposition of  $|e_i\rangle$ :  $|v\rangle = \sum_i v_i |e_i\rangle$ , where  $\sum_i v_i^2 = 1$ . The inner product of  $\vec{u}$  and  $\vec{v}$  is represented as  $\langle u|v\rangle$ . For a unit vector  $|u\rangle$ , the projector is denoted as  $|u\rangle\langle u| \in \mathbb{R}^{n \times n}$ .  $|u\rangle\langle u|$  can also represent a density matrix of a pure state. A real density matrix  $\rho$  is symmetric,  $\rho = \rho^T$ , positive semi-definite,  $\rho \geq 0$ , and of trace 1,  $tr\rho = 1$ . The set of  $n \times n$  real density matrices would be noted  $\mathbb{S}^n$ .

A quantum probability measure  $\mu$  is the generalization of a classical probability measure. It satisfies two conditions: (i) for each projector  $|v\rangle\langle v|$ ,  $\mu(|v\rangle\langle v|) \in [0, 1]$  and (ii) for any orthogonal basis  $\{|u_i\rangle\}$  for  $\mathbb{R}^n$ , we have  $\sum_{i=1}^n \mu(|u_i\rangle\langle u_i|) = 1$ . The Gleason's Theorem [24] has proven the existence of a mapping function  $\mu_\rho(|v\rangle\langle v|) = tr(\rho|v\rangle\langle v|)$  for any vector  $\vec{v}$ .

### 3.2. Two-State Vector Formalism

In the standard formalism of QT, a quantum system is described by a single forward-evolving vector

$$|\Psi\rangle \quad (1)$$

which is also named a pre-selected system, as shown in Figure1- (a). The non-degenerate operator  $A$  measuring on a given vector  $|\Psi\rangle$  yields an eigenvalue  $a_k$  with the probability:

$$Pr(a_k|\Psi) = |\langle a_k|\Psi\rangle|^2 \quad (2)$$

Under the standard formalism, the maximal information contained in such a pre-selected system at present is constrained by the measurement results in the past. However, one single vector cannot completely characterize the current quantum state [25], since the 'future' information is not taken into consideration [14]. Accordingly, Aharonov, Bergman, and Lebowitz (ABL)[10] proposed a framework named Two-State Vector Formalism (TSVF), which postulated that a more complete description of a quantum system should be given by a two-state vector:

$$\langle\Psi_{post}||\Psi_{pre}\rangle \quad (3)$$

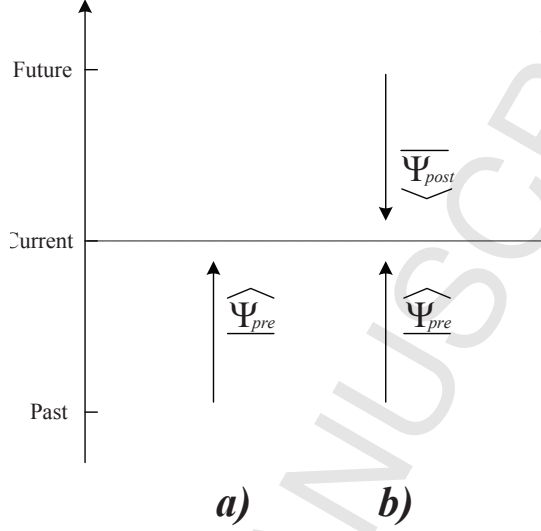


Figure 1: Description of quantum systems: (a) a pre-selected system under the standard formalism, (b) a pre- and post-selected system under Two-State Vector Formalism. Note that  $\Psi_{pre}$  is a ket  $|\Psi_{pre}\rangle$ , and  $\Psi_{post}$  is a bra  $\langle\Psi_{post}|$

where  $|\Psi_{pre}\rangle$  is a pre-selected state evolving from past to now (i.e. forward-evolving) and  $\langle\Psi_{post}|$  is a post-selected state evolving from the future to now (i.e. backward-evolving), respectively shown in Figure 1- (b). ABL also proves that an intermediate measurement of the non-degenerate operator  $B$  yields an eigenvalue  $b_k$  with the probability[12]:

$$Pr(b_k|\Psi_{pre}, \Psi_{post}) = \frac{|\langle\Psi_{post}|b_k\rangle|^2|\langle b_k|\Psi_{pre}\rangle|^2}{\sum_j |\langle\Psi_{post}|b_j\rangle|^2|\langle b_j|\Psi_{pre}\rangle|^2} \quad (4a)$$

$$= \frac{|\langle\Psi_{post}|b_k\rangle\langle b_k|\Psi_{pre}\rangle|^2}{\sum_j |\langle\Psi_{post}|b_j\rangle\langle b_j|\Psi_{pre}\rangle|^2} \quad (4b)$$

$$= \frac{|\langle\Psi_{post}|P_{B=b_k}|\Psi_{pre}\rangle|^2}{\sum_j |\langle\Psi_{post}|P_{B=b_j}|\Psi_{pre}\rangle|^2} \quad (4c)$$

where  $P_{B=b_k} = |b_k\rangle\langle b_k|$  is the projector of state  $b_k$  and Eq(4c) is named *ABL rule*. Note that if summing over a complete set of the post-selected states, the *ABL rule* will be transformed into the regular probability formalism as Eq(2). The left hand of the Eq(4) is a conditional probability which has an expansion as following:

$$Pr(b_k|\Psi_{pre}, \Psi_{post}) = \frac{Pr(b_k, \Psi_{post}|\Psi_{pre})}{Pr(\Psi_{post}|\Psi_{pre})}, \quad (5)$$

which is expanded by Eq(4) according to Bayesian Theory. And the denominator of Eq(4c) is a normalization factor. So, Eq(5) can be approximated as following:

$$Pr(b_k|\Psi_{pre}, \Psi_{post}) \propto Pr(b_k, \Psi_{post}|\Psi_{pre}) \quad (6a)$$

$$= |\langle \Psi_{post} | b_k \rangle|^2 |\langle b_k | \Psi_{pre} \rangle|^2 \quad (6b)$$

$$= Pr(\Psi_{post}|b_k) Pr(b_k|\Psi_{pre}) \quad (6c)$$

From the Eq(6c), the joint probability of the intermediate event can be approximately equaled to the product of the probability of the pre-selected state and the probability of the post-selected state. This equation also makes an inspiration for us that when estimating the probability distribution of a retrieved document in session search; both users' 'history' information needs and 'future' information needs should be considered. In the next section, we will make a discussion about how to represent users' 'history' and 'future' information needs under the framework of TSVF.

From a simple example, we may clarify the different description of quantum system between standard QT and TSVF. Let us consider a familiar cognition scenario where Alice is required to make an appraisalment for Bob. Intuitively, Bob may obtain more accurate and pertinent appraisalment if Alice has known Bob deeply than that when they just met. A possible reason is that the maximal information they can exploit in the Einsten-Podolsky-Rosen sense is limited when they just met. While, when they met after months (can be regarded as the 'future' compared with the start), the maximal information they can exploit will be enlarged significantly, thus led to a deeper mutual understanding.

### 3.3. Analogy to TSVF in IR

In this section, enlightened with the Two-State Vector Formalism (TSVF), the quasi-current IN is constructed, and the really necessary information is also filtered for modeling session search task. This paper uses a quantum system as an analogy of users' IN. Under the standard formalism of QT, users' IN should be a forward-evolving system, as shown in Figure2- (a). And it is currently and mainly inferred by users' interaction information in search engine, such as existing queries, clicked documents and skipped documents, etc. However, there is another property of QT that the evolution of quantum system is a Markov-like process that the maximal information contained in the current quantum

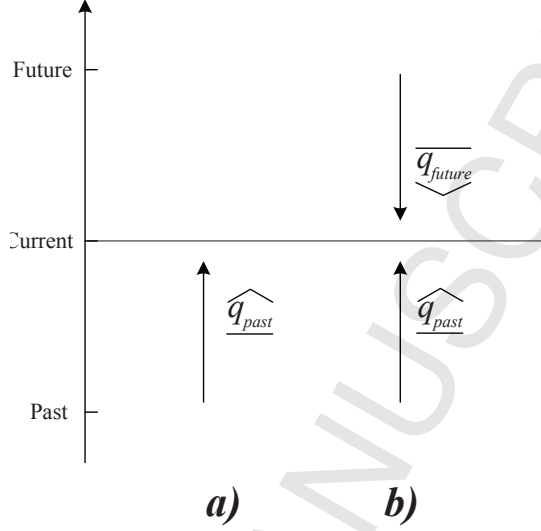


Figure 2: Description of quantum-based IR system: (a) existing quantum-based IR models (b) a quantum-based IR model inspired by TSVF. Note that  $q_{past}$  is a ket  $|q_{past}\rangle$ , and  $q_{future}$  is a bra  $\langle q_{future}|$

state is determined by the results of the last measurement. According to this, users' current IN should be extracted from the current query, and the historical interaction data should have no influence on the current IN. Besides, such a representation is not complete by limiting the current state on the last historical measurement results, and the information from future is not taken into consideration. To seek a more complete description, we need to find necessary information both from future and past to approach to describe users' current IN.

In this paper, the quasi-current IN is assumed to be derived from the nearest future query and the last historical query. Our idea is in accordance with the Figure2- (b), which the future IN corresponds to the vector  $q_{future}$  and the historical IN corresponds to the vector  $q_{past}$ . The nearest future query is actually the current query that has not been retrieved by search engine yet. It is the most relevant information with respect to users' really current information needs, and represents the prediction of really relevant information as well. What's more, the real future information is implicit and cannot be captured directly. Therefore, it is reasonable to treat the current query as the nearest future query. The last historical query is the query before the current query, also named the previous

query. Intuitively, it is the closest historical query for representing the current IN. During a session search, with the change of users' IN, modified queries are getting closer to the real relevant information, and the initial query deviates much more from current IN in the meantime. Especially for a multi-query task, the initial query is greater far away from users' real IN than the current query. Thus, if those earlier information was applied for IR tasks, they may disturb the search engine to capture the real IN. To avoid redundant interactive information and utilize it properly, the previous query is used to replace all the historical interaction data.

We take an example from Session Track data to show the reason why the TSVF framework is adopted in session-based IR settings. And we will make a further analysis based on Quantum Language Model (QLM) which uses the density matrices to represent the probability distribution of the documents. The participant in TREC 2013 Session 3 writes three queries described as follow:

- query 1: "heart attack causes"
- query 2: "heart attack causes nhs"
- query 3: "heart attack causes site:nhs.uk"

According to TSVF, query 2 and query 3 will be sent to search engine to represent users' current IN. It is apparently reasonable that both queries describe similar IN and share the same key words "heart attack causes nhs". When estimating the density matrix of users' current IN (two queries), the probability of these key words will be increased. And the score of estimated documents containing the same key words will be increased as well. If query 1 is added to represent users' IN, the probability of "heart attack causes" will be increased. The evaluated documents mainly containing "heart attack causes" will be ranked at the top of returned list. However, it is actually derived from users' real IN that they want to find a website relevant with "heart attack causes". **Therefore, compared with just utilizing current query or all of historical queries, a two-state vector (extracted from the current query and the previous query in our tasks) provides a more complete and proper description for the quasi-current information needs.**

#### 4. A Quantum IR Model inspired by TSVF

In order to realize our idea, we stem from the computational framework proposed by a recent Quantum Language Model (QLM) approach for IR [5].  
 255 As an extension for classical unigram Language Model, QLM can be used to capture richer information than single terms from text excerpts. Document or query is represented by a density matrix, a well-known mathematical object in physics. The quantum relative entropy is applied to score the similarity between evaluated documents and a given query. Our contribution here is to show that  
 260 our quasi-current IN extracted from both the current query and the previous query is effective for session search task. From now on, we will introduce our quantum-based IR model inspired by TSVF (**QMT**).

##### 4.1. Representation

In line with the original QLM approach, each single term or compound dependency in the vocabulary is expressed as a projector  $\Pi_k$ . For the query  $Q$ , supposing the set  $|e_{w_j}\rangle$  forms an orthogonal basis. The projector for a single word  $w_j$  is below:

$$m(w_j) = |e_{w_j}\rangle\langle e_{w_j}|, w_j \in Q. \quad (7)$$

And the projector for a compound dependency  $k$  (with two or more words for each dependency) is below:

$$m(k) = m(\{w_1, \dots, w_k\}) = |k\rangle\langle k|, |k\rangle = \sum_{i=1}^K \delta_i |e_{w_i}\rangle \quad (8)$$

270 where the coefficients  $\delta_i \in \mathbb{R}$  must be chosen such that  $\sum_i \delta_i^2 = 1$ , and can be assigned to the uniform ( $\delta_i = \sqrt{1/n}$ ) or the normalized inverse document frequency (*idf*) weight. In the original QLM, the current IN is described by a single query. According to our analysis above, our proposed quasi-current IN should be described by two queries, the nearest future query (the current query)  
 275  $Q_{future}$ , and the last historical query (the previous query)  $Q_{past}$ . In our model (QMT), the single words, bi-grams and tri-grams are considered as possible compound dependencies that are extracted from those two queries. The past projector set is denoted as  $\mathcal{P}_{past} = \{\Pi_i\}_{i=1}^{M_{past}}$  and the future projector set is denoted as  $\mathcal{P}_{future} = \{\Pi_j\}_{j=1}^{M_{future}}$ , where  $M$  is the number of projectors. Then,  
 280 for each occurrence of single terms and compound dependencies in unordered fixed windows (with length  $L$ ) of a document, those projectors are added to the

sequence of projectors for the document. We choose to parameterize  $\#$  as the unordered window operator in **Indri** ( $\#uwL$ ), an open source search engine. The Algorithm 1 shows the process of building the sequence of the future and past projector sets for a document  $d$ .

---

**Algorithm 1** builds the sequence  $\mathcal{P}_{future}$  and  $\mathcal{P}_{past}$  for a document  $d$  given  $Q_{future}$  and  $Q_{past}$

---

**Require:**  $Q_{future}, Q_{past}$

```

1:  $\mathcal{P}_{future} \leftarrow \phi, \mathcal{P}_{past} \leftarrow \phi;$ 
2: // extract projectors from  $Q_{future}$ 
3: for each  $k \in \mathcal{P}(Q_{future})$  do
4:   //  $k$  is a single term or a compound dependency
5:   for  $i = 1; i \leq \#(k, d); i++$  do
6:     // add the projector to the sequence
7:      $\mathcal{P}_{future} \leftarrow \mathcal{P}_{future} \oplus m(k);$ 
8:   end for
9: end for
10: // extract projectors from  $Q_{past}$ 
11: for each  $k \in \mathcal{P}(Q_{past})$  do
12:   //  $k$  is a single term or a compound dependency
13:   for  $i = 1; i \leq \#(k, d); i++$  do
14:     // add the projector to the sequence
15:      $\mathcal{P}_{past} \leftarrow \mathcal{P}_{past} \oplus m(k);$ 
16:   end for
17: end for
18: return  $\mathcal{P}_{future}, \mathcal{P}_{past}$ 

```

---

285

After obtaining the sequence of projectors from the evaluated document, we define the target function for training the density matrix  $\rho$ . In the original QLM, given a sequence of projectors  $\mathcal{P}_d = \{\Pi_i : i = 1, \dots, M\}$  for a document  $d$ , the estimation of density matrix  $\rho_d$  can be transformed to the following maximization problem, in which the objective function is the logarithm of the likelihood:

290

$$\max_{\rho_d} \log \prod_{i=1}^M \text{tr}(\rho_d \Pi_i) \quad (9)$$

where each  $\Pi_i$  is a quantum elementary event representing a single term or

compound dependency, and  $M$  is the number of quantum events (projectors).

Distinct with QLM, we propose a more general framework for session search task. According to the analysis above, a more complete description for the current information needs should be composed both by its past (the previous query) and future (the current query). The Eq(6c) provides a mathematical framework for estimating the probability distribution (density matrix) of the retrieved documents. Therefore, the current likelihood function of the estimated document should be the product of the likelihood of the future IN and the likelihood of the past IN. The target function in our model is formalized as follow:

$$\mathcal{L}_{current}(\rho_d) = \mathcal{L}_{\mathcal{P}_{past}}(\rho_d)\mathcal{L}_{\mathcal{P}_{future}}(\rho_d) \equiv F(\rho_d). \quad (10)$$

The estimated  $\rho_d$  can be obtained by approximately solving the following maximization problem:

$$\rho_d = \arg \max \left( \sum_{i=1}^{M_{past}} \log \text{tr}(\rho_d \Pi_i) + \sum_{j=1}^{M_{future}} \log \text{tr}(\rho_d \Pi_j) \right). \quad (11)$$

Note that  $M_{past}$  cannot be merged with  $M_{future}$  because  $\mathcal{P}_{past}$  and  $\mathcal{P}_{future}$  are extracted from the previous query and the current query, respectively.

#### 4.2. Learning

The target function Eq(11) of our model is similar to the objective function Eq(9) in original QLM. Thus, the  $R\rho R$  algorithm [26] can be used to solve the optimization problem. The  $R\rho R$  algorithm works in most of cases, but it has no theoretical guarantee of convergence, regardless the dataset and the initial point[17]. So, in this paper, we decide to use another algorithm [17] proposed under a line search procedure with Armijo condition, namely “Diluted  $R\rho R$ ”. Besides globally convergent, the algorithm is computationally practicable as well. During the optimization process, the search direction is a combination of two ascent directions controlled by the step size  $t$ . In the paper [17], it has shown that an inexact line search method to determine  $t$  is enough for finding a value to guarantee the global convergence. The initial density matrix  $\tilde{\rho}_d(0) = \text{diag}(\theta^{ML})$ , where  $\theta^{ML}$  is a classical maximum likelihood language model of a document or a query. The search direction is given by:

$$D^k = \left( \frac{2}{q(t_k)} \bar{D}^k + \frac{t_k \text{tr}(\nabla F(\tilde{\rho}_d(k)) \tilde{\rho}_d(k) \nabla F(\tilde{\rho}_d(k)))}{q(t_k)} \tilde{D}^k \right) \quad (12)$$



where  $\nabla F(\tilde{\rho}_d(k))$ ,  $q(t_k)$ ,  $\bar{D}^k$ ,  $\tilde{D}^k$  are listed as follow:

$$\nabla F(\tilde{\rho}_d(k)) = \sum_i \frac{f_i}{\text{tr}(\tilde{\rho}_d(k)\Pi_i)} \Pi_i, \quad f_i = \frac{M_i}{M}. \quad (13)$$

Note that  $M_i$  is the number of  $\Pi_i$ ,  $M$  is the total of all the quantum events (projectors).

$$\bar{D}^k = \frac{\nabla F(\tilde{\rho}_d(k))\tilde{\rho}_d(k) + \tilde{\rho}_d(k)\nabla F(\tilde{\rho}_d(k))}{2} - \tilde{\rho}_d(k) \quad (14)$$

$$\tilde{D}^k = \frac{\nabla F(\tilde{\rho}_d(k))\tilde{\rho}_d(k)\nabla F(\tilde{\rho}_d(k))}{\text{tr}(\nabla F(\tilde{\rho}_d(k))\tilde{\rho}_d(k)\nabla F(\tilde{\rho}_d(k)))} - \tilde{\rho}_d(k) \quad (15)$$

325

$$q(t_k) = 1 + 2t_k + t_k^2 \text{tr}(\nabla F(\tilde{\rho}_d(k))\tilde{\rho}_d(k)\nabla F(\tilde{\rho}_d(k))) \quad (16)$$

The the step size  $t_k$  is updated according to the condition described as follow:

$$F(\tilde{\rho}_d(k) + t_k \bar{D}^k) \leq F(\tilde{\rho}_d(k)) + \gamma t_k \text{tr}(\nabla F(\tilde{\rho}_d(k))\bar{D}^k) \quad (17)$$

If Eq(17) is satisfied, it is shown to be not convergent at global optimum. Then a new step size  $t_k$  less than 1 is chosen, and is provided for Eq(12) to get a new search direction  $D$ . If the condition Eq(17) is broken, the density matrix will be updated as follow:

330

$$\tilde{\rho}_d(k+1) = \tilde{\rho}_d(k) + t_k D^k \quad (18)$$

In this paper, we set the number of iterations to 100,  $\gamma$  to  $10^{-4}$ . The initial value of  $t$  is set to 1, and every time it updates with multiplying by 0.7.

#### 4.3. Smoothing

After training a density matrix, the Dirichlet smoothing method is applied to smooth the density matrices. If  $\tilde{\rho}_d$  is a document QLM obtained by MLE, its smoothed version is obtained by interpolation with the MLE collection QLM  $\tilde{\rho}_c$ :

335

$$\rho_d = (1 - \alpha_d)\tilde{\rho}_d + \alpha_d\tilde{\rho}_c \quad (19)$$

where  $\alpha_d \in [0, 1]$  controls the amount of smoothing, and  $\alpha_d = \frac{\mu}{\mu + M}$  is a well-known form of the parameter for Dirichlet smoothing.  $\mu$  is set to 5000, and  $M$  is the number of quantum events happened in the collection.

340

Table 1: Statistics for TREC 2013 and 2014 Datasets (TREC 2014’s official ground truth only contains the first 100 sessions).

Items	TREC 2013	TREC 2014
#Sessions	87	100
#Queries	442	453
#Avg.session length	5.08	4.53

#### 4.4. Scoring

After obtaining  $\rho_Q$  and  $\rho_d$ , the negative Von Neumann (VN) divergence is used as the scoring function to rank the documents:

$$-\triangle_{VN}(\rho_Q||\rho_d) = -tr(\rho_Q(\log(\rho_Q) - \log(\rho_d))) \quad (20)$$

$$\stackrel{rank}{=} tr(\rho_Q \log(\rho_d))$$

In accordance with the analysis in [5], the VN divergence can offer a way to  
 345 analyze how much the query is relevant to the document.

## 5. EMPIRICAL EVALUATION

### 5.1. Data Set

Our experiments are conducted on TREC Session tracks 2013 and 2014. Table 1 lists the statistic information about these two datasets. Session track  
 350 2014 actually owns 1024 sessions. Since the official assessors only assessed the first 100 sessions, we only take tests on this subset. For each session, there are a current query used for retrieval task, and several historical queries, search results, clicked documents, and dwell time. The corpus used in our experiments is the Clueweb12 Category B collection (CatB) which contains more than 50  
 355 million English webpages collected from the Internet. The collection is indexed with Indri 5.6, meanwhile, all words are stemmed by the Porter stemmer and stop words are removed according to the standard stop words list. In order to simulate real web search scenario and illustrate the robustness of our proposed model, no spam filtering is adopted for all the runs reported in this paper.

### 5.2. Descriptions for Tested Models

To verify the effectiveness of our proposed model, both typical IR models and Quantum-based IR (QIR) models are used to make comparisons. The tested models are described as following:

Table 2: Parameter settings

IR Models	Parameters for TREC 2013	Parameters for TREC 2014
Unigram	-	-
QCM	$\alpha_{qcm} = 2.2, \beta_{qcm} = 1.8, \epsilon_{qcm} = 0.07, \delta_{qcm} = 0.4, \gamma_{qcm} = 0.92$	
RM-HS	$fbTerm = 50, \lambda = 0.7$	
QLM	$L = 16, \beta = 0.2$	$L = 8, \beta = 0.8$
QQE	$L = 32, \beta = 0.2$	$L = 8, \beta = 0.8$
SQLM	$L = 16, \beta = 0.2$	$L = 4, \beta = 0.5$
QMT	$L = 6, \beta = 0.2$	$L = 1, \beta = 0.1$

1. Unigram, a classical unigram Language Model as our baseline model.
- 365 2. QCM, a Query Change Model [27], which utilizes query change and previously retrieved documents to enhance session search.
3. RM-HS, a query expansion approach, in which pseudo feedback documents are replaced by historical queries and clicked snippets[8].
4. QLM, a Quantum Language Model [5], which adopts the quantum probabilistic framework to model term dependencies for text excerpts.
- 370 5. QQE, a quantum-based IR model with quantum entanglement (QE) [23], which proves the statistic connection between QE and Unconditional Pure Dependence, and is proposed based on QLM.
6. SQLM, a session-based QLM [7], which uses the transformation between density matrices to model the evolution of users' information needs.
- 375 7. QMT, the proposed QIR Model in this paper inspired by TSVF.

Note that, Unigram, QLM and QQE use the last query. QCM, RM-HS and SQLM exploit all the queries and document information in the session. According to analysis above, our proposed model (QMT) utilizes the last two queries (current and previous query).

380

For all the IR models, the Dirichlet smoothing parameter  $\mu$  is set to 5000. And these models are evaluated under the same ground truth and the same evaluation metrics. The top ranked 1000 documents retrieved with the unigram Language Model are used to re-rank for all tested models. Since the top ten retrieved documents are the most concerned results for users, we employ the nDCG@10 (Discount Cumulative Gain) and MAP@10 (Mean average precision) to evaluate those tested models. NDCG@k is the official evaluation metric

385

[28, 29] and can evaluate the relevance degree of documents. MAP@k can reflect the precision of top  $k$  retrieved results and is also another important metric for session search task [29]. Further, nDCG@1 and nDCG@5 are also used to prove the effectiveness of our model. Note that, the final score of all QLM-based models are combined with first round results by a linear parameter  $\beta$ ,  $\beta \in (0, 1)$  with the increment 0.1:

$$rScore(d) = (1 - \beta) \times oScore(d) + \beta \times qScore(d) \quad (21)$$

where  $d$  is a document,  $oScore$  is the original score of  $d$ , and  $qScore$  is the score obtained by QIR models for  $d$ . We use uniform superposition weights (i.e.  $\delta_i = \sqrt{1/n}$ ) for quantum event in QLM, and test different fixed-window sizes (i.e. different  $L \in \{1, 2, 4, 6, 8, 16, 32\}$ ). More detailed parameter settings with optimal performance are listed in Table 2.  $\alpha_{qcm}, \beta_{qcm}, \epsilon_{qcm}, \delta_{qcm}, \gamma_{qcm}$  are the original optimal parameters in QCM. We only report the metric nDCG@10 of RM-HS in this paper due to the following reasons: first, it follows the same experimental environment with our experiments, thus, we do not realize this model; second, the original paper [8] about RM-HS only mentioned the metric NDCG@10, and other metrics values in this paper are not computed.  $fbTerm$  in RM-HS is the number of pseudo feedback terms, and  $\lambda$  is the weight of expanded query terms.

### 5.3. Results and Discussion

In this section, we report and analyze the performance for all the tested models on Session Track 2013 and 2014. The evaluation results for all models evaluated with nDCG@1, nDCG@5, nDCG@10 and MAP @10 are reported in Table 3 and Table 4. From the tables, we can find that all QLM-based IR models (i.e. QLM, QQE, SQLM, QMT) enhance the retrieval effectiveness and stability in comparison with the classical IR models (i.e. Unigram, QCM, RM-HS) across two data sets with respect to all the evaluation metrics. The good performance can be attributed to the reason that QLM can model term dependencies and capture richer information than single terms from text excerpts. Then, we further make comparison among QLM-based IR models and find that SQLM using interaction information of session does not outperform QLM which only use single current query. It shows that not all the interaction data is necessary for modeling session search.

Table 3: Performance on TREC 2013. Significance test has been conducted for all expansion models compared with Unigram with paired  $t$ -test. Symbol  $\alpha$  means  $p < 0.01$ ,  $\beta$  means  $p < 0.05$ , and boldface means the best performance.

IR Models	TREC 2013			
	nDCG@1	nDCG@5	nDCG@10	MAP@10
Unigram	0.0565 (0.00)	0.0558 (0.00)	0.0605 (0.00)	0.0072 (0.00)
QCM	0.0672 (18.99 $^{\beta}$ )	0.0597 (7.01)	0.0613 (1.45)	0.0110 (53.63 $^{\alpha}$ )
RM-HS	-	-	0.0600 (-0.82)	-
QLM	0.0849 (50.39 $^{\alpha}$ )	0.0673 (20.52 $^{\alpha}$ )	0.0670 (10.77 $^{\beta}$ )	0.0090 (24.65 $^{\alpha}$ )
QQE	0.0795 (40.70 $^{\alpha}$ )	0.0635 (13.86 $^{\beta}$ )	0.0634 (4.81)	0.0078 (8.10)
SQLM	0.0664 (17.64 $^{\beta}$ )	0.0558 (0.00)	0.0605 (0.06)	0.0072 (0.04)
QMT	<b>0.1337</b> (136.63 $^{\alpha}$ )	<b>0.0982</b> (75.98 $^{\alpha}$ )	<b>0.0888</b> (46.77 $^{\alpha}$ )	<b>0.0123</b> (70.83 $^{\alpha}$ )

Table 4: Performance on TREC 2014. Significance test has been conducted for all expansion models compared with Unigram with paired  $t$ -test. Symbol  $\alpha$  means  $p < 0.01$ ,  $\beta$  means  $p < 0.05$ , and boldface means the best performance.

IR Models	TREC 2014			
	nDCG@1	nDCG@5	nDCG@10	MAP@10
Unigram	0.1427 (0.00)	0.1396 (0.00)	0.1422 (0.00)	0.0215 (0.00)
QCM	0.0961 (-32.64)	0.1229 (-11.96)	0.1302 (-8.40)	0.0174 (-19.16)
RM-HS	-	-	0.1393 (-0.29)	-
QLM	0.1536 (7.68)	0.1422 (1.81)	0.1427 (0.38)	0.0215 (0.34)
QQE	0.1641 (15.02 $^{\beta}$ )	0.1429 (2.33)	0.1428 (0.48)	<b>0.0216 (0.78)</b>
SQLM	0.1507 (5.61)	0.1422 (1.86)	0.1444 (1.58)	0.0215 (0.35)
QMT	<b>0.1650</b> (15.62 $^{\alpha}$ )	<b>0.1529</b> (9.52 $^{\beta}$ )	0.1489 (4.71 $^{\beta}$ )	<b>0.0194 (-9.76)</b>

Our model, only using current query and the previous query, can mostly achieve the best performance over either classical IR models or QLM-based IR models. Especially for the evaluation metric nDCG, which can evaluate the ranking of relevant documents, our proposed QMT achieves significantly improvement over other tested model. Despite our model does not make improvement on TREC 2014 with respect to MAP@10, it still can push relevant contents to the front of search results other than just retrieve them. Beyond that, session length (the number of interactions in a session) is also a key factor that can influence the performance of the different models on different data sets. The longer the session length, the richer the interaction information involved in that session. And the retrieval area of two adjacent queries will become similar as well. From Table 1, the average session length of TREC 2013 is longer than TREC 2014, which means that the last two queries of the former pool more similar retrieval information than that of the latter. The distinction between the previous query and the current query in TREC 2014 may bring more noise information than TREC 2013 during a search, and further lower the precision of retrieved results.

## 6. Conclusions and Future Work

In this paper, we present an innovative analogy between an emerging Two-State Vector Formalism (TSVF) theory and a dynamic IR task. From the analogy, we construct a quasi-current IN and find that a “two-state vector” (two nearest queries in our tasks) can provide necessary information for modeling session search task. Moreover, based on analysis from quantum theory, the previous query can replace all the interaction information in the session of search engine, and the current query can be regarded as the nearest future information. Inspired by the TSVF, we propose a novel quantum information retrieval model (QMT) which can simulate users’ TSVF-like cognition process in web search. Extensive experiments have been conducted on the session tracks of TREC 2013 and 2014, which show that QMT outperforms a series of compared IR models.

In the future, we can further improve our model by investigating more effective means (e.g., pseudo-relevance feedback, EEG and eye tracker) to estimate the future information and extract precisely history information from user behaviors (click, skip and dwell). Moreover, TSVF may also motivate the

development of other research problems, such as the time-sensitive prediction tasks and context-sensitive intention understanding problem.

## References

- [1] C. J. Van Rijsbergen, The geometry of information retrieval, Cambridge University Press, 2004.
- [2] B. Piwowarski, I. Frommholz, M. Lalmas, K. Van Rijsbergen, What can quantum theory bring to information retrieval, in: Proceedings of the 19th ACM international conference on Information and knowledge management, ACM, 2010, pp. 59–68.
- [3] X. Zhao, P. Zhang, D. Song, Y. Hou, A novel re-ranking approach inspired by quantum measurement, in: European Conference on Information Retrieval, Springer, 2011, pp. 721–724.
- [4] G. Zuccon, L. Azzopardi, Using the quantum probability ranking principle to rank interdependent documents, in: European Conference on Information Retrieval, Springer, 2010, pp. 357–369.
- [5] A. Sordoni, J.-Y. Nie, Y. Bengio, Modeling term dependencies with quantum language models for ir, in: Proceedings of the 36th international ACM SIGIR conference on Research and development in information retrieval, ACM, 2013, pp. 653–662.
- [6] J. Li, P. Zhang, D. Song, Y. Hou, An adaptive contextual quantum language model, Physica A: Statistical Mechanics and its Applications 456 (2016) 51–67.
- [7] Q. Li, J. Li, P. Zhang, D. Song, Modeling multi-query retrieval tasks using density matrix transformation, in: Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval, ACM, 2015, pp. 871–874.
- [8] P. Zhang, J. Li, B. Wang, X. Zhao, D. Song, Y. Hou, M. Melucci, A quantum query expansion approach for session search, Entropy 18 (4) (2016) 146.
- [9] Y. Aharonov, L. Vaidman, The two-state vector formalism: an updated review, in: Time in quantum mechanics, Springer, 2008, pp. 399–447.

- [10] Y. Aharonov, P. G. Bergmann, J. L. Lebowitz, Time symmetry in the quantum process of measurement, *Physical Review* 134 (6B) (1964) B1410.
- 485 [11] T. Ravon, L. Vaidman, The three-box paradox revisited, *Journal of Physics A: Mathematical and Theoretical* 40 (11) (2007) 2873.
- [12] Y. Aharonov, L. Vaidman, Complete description of a quantum system at a given time, *Journal of Physics A: Mathematical and General* 24 (10) (1991) 2315.
- 490 [13] B. Gibran, Causal realism in the philosophy of mind, *Essays in Philosophy* 15 (2) (2014) 5.
- [14] L. Vaidman, Y. Aharonov, D. Z. Albert, How to ascertain the values of  $\sigma_x$ ,  $\sigma_y$ , and  $\sigma_z$  of a spin-1/2 particle, *Physical review letters* 58 (14) (1987) 1385.
- 495 [15] Y. Aharonov, D. Rohrlich, *Quantum paradoxes: quantum theory for the perplexed*, John Wiley & Sons, 2008.
- [16] Y. Aharonov, S. Popescu, D. Rohrlich, P. Skrzypczyk, Quantum cheshire cats, *New Journal of Physics* 15 (11) (2013) 113015.
- [17] D. S. Goncalves, M. A. Gomes-Ruggiero, C. Lavor, Global convergence of diluted iterations in maximum-likelihood quantum tomography, *arXiv preprint arXiv:1306.3057*.
- 500 [18] P. Zhang, D. Song, X. Zhao, Y. Hou, Investigating query-drift problem from a novel perspective of photon polarization, in: *Conference on the Theory of Information Retrieval*, Springer, 2011, pp. 332–336.
- 505 [19] A. Sordoni, J. He, J.-Y. Nie, Modeling latent topic interactions using quantum interference for information retrieval, in: *Proceedings of the 22nd ACM international conference on Information & Knowledge Management*, ACM, 2013, pp. 1197–1200.
- 510 [20] M. Melucci, An investigation of quantum interference in information retrieval, in: *Information Retrieval Facility Conference*, Springer, 2010, pp. 136–151.



- [21] B. Wang, P. Zhang, J. Li, D. Song, Y. Hou, Z. Shang, Exploration of quantum interference in document relevance judgement discrepancy, *Entropy* 18 (4) (2016) 144.
- 515 [22] A. Sordoni, Y. Bengio, J.-Y. Nie, Learning concept embeddings for query expansion by quantum entropy minimization., in: *AAAI*, Vol. 14, 2014, pp. 1586–1592.
- [23] M. Xie, Y. Hou, P. Zhang, J. Li, W. Li, D. Song, Modeling quantum entanglements in quantum language models.
- 520 [24] A. M. Gleason, Measures on the closed subspaces of a hilbert space, in: *The Logico-Algebraic Approach to Quantum Mechanics*, Springer, 1975, pp. 123–133.
- [25] S. Watanabe, Symmetry of physical laws. part iii. prediction and retrodiction, *Reviews of Modern Physics* 27 (2) (1955) 179.
- 525 [26] A. Lvovsky, Iterative maximum-likelihood reconstruction in quantum homodyne tomography, *Journal of Optics B: Quantum and Semiclassical Optics* 6 (6) (2004) S556.
- [27] D. Guan, S. Zhang, H. Yang, Utilizing query change for session search, in: *Proceedings of the 36th international ACM SIGIR conference on Research and development in information retrieval*, ACM, 2013, pp. 453–462.
- 530 [28] B. Bushee, D. Pintus, P. Smith, S. G. Small, Sira: Trec session track 2014, Tech. rep., DTIC Document (2014).
- [29] B. Carterette, E. Kanoulas, M. Hall, P. Clough, Overview of the trec 2014 session track, Tech. rep., DTIC Document (2014).